ON AN APPLICATION OF MARKOV CHAIN TO NUMERICAL ANALYSIS

MASAHIKO SHINOHARA

(Received October 26, 1988)

Summary

We treat a special type of Markov chain with a finite state space. This type of Markov chain often appears in traffic theory, especially in the theoretical model of the multi-channel telecommunication via some electronic device, say satellite, for which there is an upper limit of the amount of information transmitted per unit time. In such an application we are interested in its stationary state, that is the invariant measure of the Markov chain. We propose an iteration procedure to compute the invariant measure numerically by a digital computer. The original Markov chain has a continuous time parameter. Our recipe is to introduce another Markov chain having the same invariant measure but with a discrete time parameter. It can be known that for any initial distribution the distribution of the latter chain converges to its invariant measure, therefore iteration of multiplying the transition probability matrix to some initial distribution vector is enough to get the invariant measure. This can be easily undergone numerically by a digital computer. Compared with the conventional method to solve a system of linear equations, our method uses far less working area of the computer memory and this is a great advantage, especially in the case when the number of the states is very large.

1. Model

We are thinking of a telecommunication device with the maximum information transmission rate of n bits/sec. This device is used for multi-channel service. More exactly, there are l kinds of services and for one service of the ith kind, the information transmission rate of m_i bits/sec is necessary ($i = 1, \cdots, l$). Let s_i be the number of the services of the ith kind ($i = 1, \cdots, l$), then the total amount of the information transmission rate is $\sum_{i=1}^{l} m_i s_i$ bits/sec, which should be equal or less than n bits/sec, according to the capacity of the device. Therefore the set S of all possible states of the communication via this

* Department of Mathematics, Tokyo Woman's Christian University, Zenpukuji, Suginami-ku, Tokyo 167, Japan.

device is the following set of l-dimentional lattice points.

$$S = \{\boldsymbol{s} = (s_1, s_2, \cdots, s_l) \mid \sum_{i=1}^{l} m_i s_i \leqq n, \text{ each } s_i \text{ is a nonnegative integer}\}$$

Next, we suppose that the calls and the terminations of the services occur independently of each other. Moreover we suppose that mean frequency of the calls for the ith service is a_i/h_i times/sec, and that its service time is distributed exponentially with mean h_i sec, where a_i and h_i are positive numbers.

Under these set-ups as stated above, we think of the telecommunication process via the device, as a Markov chain on the state space S. This is a birth and death process on the finite state space S, and the forward equations for its distribution $\{p_t(\boldsymbol{s})\}_{\boldsymbol{s} \in S}$ at time t are the following,

$$\left.\begin{array}{c} \dfrac{d}{dt} p_t(\boldsymbol{s}) = \sum_{i=1}^{l} \left\{ \dfrac{a_i}{h_i} p_t(\boldsymbol{s}^-_{(i)}) + \dfrac{s_i+1}{h_i} p_t(\boldsymbol{s}^+_{(i)}) - \left(\dfrac{a_i}{h_i}\alpha_i(\boldsymbol{s}) + \dfrac{s_i}{h_i}\right) p_t(\boldsymbol{s}) \right\} \\ \boldsymbol{s} = (s_1, s_2, \cdots, s_l) \in S \end{array}\right\}$$

where $\boldsymbol{s}^{\pm}_{(i)} = (s_1, \cdots, s_{i-1}, s_i \pm 1, s_{i+1}, \cdots, s_l)$,

$$\alpha_i(\boldsymbol{s}) = \begin{cases} 0 & \text{if } \boldsymbol{s}^+_{(i)} \notin S \\ 1 & \text{if } \boldsymbol{s}^+_{(i)} \in S \end{cases}$$

and $p_t(\boldsymbol{s}^+_{(i)})$ or $p_t(\boldsymbol{s}^-_{(i)})$ is understood to be zero if $\boldsymbol{s}^+_{(i)}$ or $\boldsymbol{s}^-_{(i)}$ does not belong to S, respectively. Let us denote the invariant measure of this process by $\{p(\boldsymbol{s})\}_{\boldsymbol{s} \in S}$, the existence and the uniqueness of which can be easily seen.

Putting the left side of the above equations 0, we get easily the following equations for $\{p(\boldsymbol{s})\}$

$$\left.\begin{array}{c} \sum_{i=1}^{l} \left\{ \left(\dfrac{a_i}{h_i}\alpha_i(\boldsymbol{s}) + \dfrac{s_i}{h_i}\right) p(\boldsymbol{s}) \right\} = \sum_{i=1}^{l} \left\{ \dfrac{a_i}{h_i} p(\boldsymbol{s}^-_{(i)}) + \dfrac{s_i+1}{h_i} (\boldsymbol{s}^+_{(i)}) \right\} \\ \boldsymbol{s} \in S \end{array}\right\}. \quad (1)$$

The solution of this system of equations is explicity written down as follows,

$$p(\boldsymbol{s}) = p(s_1, \cdots, s_l) = c \cdot \frac{a_1^{s_1} a_2^{s_2} \cdots a_l^{s_l}}{s_1! s_2! \cdots s_l!},$$

where c is the normalizing constant to satisfy $\sum_{\boldsymbol{s} \in S} p(\boldsymbol{s}) = 1$.

Therefore there is no problem left about the solution of (1), that is the invariant measure of the original Markov chain.

$$\text{Put} \quad S_i = \{\boldsymbol{s} \in S \mid \boldsymbol{s}^+_{(i)} \notin S\} \quad i = 1, \cdots, l.$$

If the device is in some state belonging S_i, then a new call for the ith service should be rejected, according to the capacity limit of the device. This may happen for each of l kinds of services. In order to improve the total efficiency of the transimission rate, several methods are proposed.

One of the methods is to reject partly the calls near S_i for the services of which m_i's are small, in order to prepare for a new call for the services of which m_i's are large. If we adopt this strategy, the equations to determine the invariant measure change from (1) to the following,

$$\left.\begin{aligned}&\sum_{i=1}^{l}\left\{\left(\frac{a_i}{h_i}\alpha_i(\boldsymbol{s})\beta_i(\boldsymbol{s})+\frac{s_i}{h_i}\right)p(\boldsymbol{s})\right\}\\&\qquad=\sum_{i=1}^{l}\left\{\frac{a_i}{h_i}\alpha_i(\boldsymbol{s}_{(i)}^{-})\beta_i(\boldsymbol{s}_{(i)}^{-})p(\boldsymbol{s}_{(i)}^{-})+\frac{s_i+1}{h_i}p(\boldsymbol{s}_{(i)}^{+})\right\}\\&\qquad\qquad\boldsymbol{s}\in S\end{aligned}\right\}, \tag{2}$$

where $\beta_i(\boldsymbol{s})$ is the rejection rate for the ith service at the state $\boldsymbol{s}$, taking the value between 0 and 1. $\beta_i=0$ means complete rejection, while $\beta_i=1$ means no rejection. $0<\beta_i<1$ means partial rejection. The solution of (2) cannot be expressed explicity as in the case of (1), and our aim is to find a simple procedure to calculate the solution of (2) numerically.

2. Iteration Method

It is not hard to see that (2) has a unique solution by the general theory of Markov chain. Let N be the number of the states belonging to S. Then (2) is a system of N linear equations, and to compute the solution of it numerically by a conventional method, the number of the computer memories which are necessary as a working area is of order N^2. This will be a serious problem when N is large. What we are now proposing is an alternative method by iteration based on the convergence theorem of Markov chain.

In order to use the matrix expression, we first line up the states of S in some linear order and give them the numbers from 1 to N. Changing the notation $\{p(\boldsymbol{s})\}_{\boldsymbol{s}\in S}$ to $\{p_i\}_{i=1,\cdots,N}$ equations (2) are expressed as follows,

$$d_k p_k=\sum_{j=1}^{N}p_j b_{jk}\qquad(k=1,\cdots,N)\,, \tag{3}$$

where $\{d_j\}_{j=1,\cdots,N}$ and $\{b_{jk}\}_{\substack{j=1,\cdots,N\\ j=1,\cdots,N}}$ are the correspondig coefficients appearing in the equations (2). All d_j's are positive and b_{jk}'s are nonnegative numbers. It is then easily seen that these coefficients satisfy the following relations,

$$d_j = \sum_{k=1}^{N} b_{jk} \qquad (j = 1, \cdots, N) . \tag{4}$$

(4) is equivalent to the condition that the Markov chain is conservative. Introducing two $N \times N$ matrices and a N-dimentional vector as

$$D = \begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_N \end{pmatrix}, \qquad B = \begin{pmatrix} b_{11} & \cdots\cdots & b_{1N} \\ \vdots & & \vdots \\ b_{N1} & \cdots\cdots & b_{NN} \end{pmatrix}$$

$$\boldsymbol{p} = \begin{pmatrix} p_1 \\ \vdots \\ p_N \end{pmatrix} .$$

equations (3) are written in the following compact form,

$${}^t\boldsymbol{p}D = {}^t\boldsymbol{p}B . \tag{5}$$

Adding $\lambda {}^t\boldsymbol{p}D$ to both sides of (5), where λ is any positive number, we get an equivalent equation,

$${}^t\boldsymbol{p}(1 + \lambda)D = {}^t\boldsymbol{p}(B + \lambda D)$$

which is also equivalent to

$${}^t\boldsymbol{p}(1 + \lambda)D = {}^t\boldsymbol{p}\{(1 + \lambda)D\}\{(1 + \lambda)D\}^{-1}(B + \lambda D) . \tag{6}$$

Put $\tilde{B} = \{(1 + \lambda)D\}^{-1}(B + \lambda D)$, and denote the entries of this matrix as $\tilde{b}_{jk}$. Then from the relations (4), $\tilde{b}_{jk}$'s satisfy the following conditions,

$$\left.\begin{array}{l} \tilde{b}_{jk} \geqq 0 \\ \sum_{k=1}^{N} \tilde{b}_{jk} = 1 \quad (j = 1, \cdots, N) \end{array}\right\} . \tag{7}$$

This means that $\tilde{B}$ is a probability matrix. Next we put

$$\tilde{\boldsymbol{p}} = (1 + \lambda)D\boldsymbol{p}$$

and denote the entries of $\tilde{\boldsymbol{p}}$ as $\tilde{p}_i$ $(i = 1, \cdots, N)$, then (6) is expressed as

$${}^t\tilde{\boldsymbol{p}} = {}^t\tilde{\boldsymbol{p}}\tilde{B} , \tag{8}$$

which means that if we consider a new Markov chain on the same state space S with discrete time parameter and having $\tilde{B}$ as its transition matrix, then $\tilde{\boldsymbol{p}}$ is its invariant measure. From the definition of the original Markov chain, it is not difficult to see that the new chain is *aperiodic and has only one class of equivalent recurrent states.* Therefore, by the general theory of Markov chains, it is concluded that starting from any initial distribution, its distribution converges to its unique invariant measure. This means that for any probability vector $\tilde{\boldsymbol{q}}$ of size N, ${}^t\tilde{\boldsymbol{q}}\tilde{B}^n$ coverges to ${}^t\tilde{\boldsymbol{p}}$ as $n \to \infty$, where $\tilde{\boldsymbol{p}}$ is the invariant measure of the new chain. This limiting procedure can be done numerically by the iteration of multiplying $\tilde{B}$ successively from the right side. Translating this proedure to that of $\boldsymbol{p}$, that is the original invariant measure, we finally get the following iteration to compute the invariant measure of the original Markov chain,

$$ {}^t\boldsymbol{p}^{(n+1)} = {}^t\boldsymbol{p}^{(n)}(B + \lambda D)\cdot\{(1+\lambda)D\}^{-1} , $$

where $\boldsymbol{p}^{(n)}$ means the vector of the nth step of the iteration.

We have to note that the sum $\sum_{i=1}^{N} d_i p_i^{(n)}$ is conserved by this iteration, but not the sum $\sum_{i=1}^{N} p_i^{(n)}$. Therefore we have to normalize the vector after convergence to get the invariant probability measure.

This method is clearly applicable to other Markov chains which satisfy similar conditions as stated above.

The choice of λ is arbitrary, but the dynamical change of λ will be possible to get speedy convergence.